Séance 3 : Comp OT

Kantorovitch : $\alpha = \sum_{i=1}^{n} a_i \delta_{x_i}$ $\beta = \sum_{j=1}^{m} b_j \cdot \delta_{y_j}$ (*)

$$\min_{P \in \mathbb{R}_+^{n \times m}} \left\{ \langle C, P \rangle : P \mathbb{1}_m = a \ , \ P^T \mathbb{1}_n = b \right\}$$

Def : $\alpha \in \mathcal{P}(X)$ $\beta \in \mathcal{P}(Y)$ A coupling $\pi \in \mathcal{P}(X,Y)$ having "marginals" α and β

$$\text{“} \int_Y d\pi(x,y) = d\alpha(x) \text{”}$$

$$\forall f \in \mathcal{C}(X), \ \iint f(x)\, d\pi(x,y) \overset{①}{=} \int f(x)\, d\alpha(x)$$

$$\forall g \in \mathcal{C}(Y), \ \iint g(y)\, d\pi(x,y) \overset{②}{=} \int g(y)\, d\beta(y)$$

Def : $P_1 : (x,y) \in X \times Y \longmapsto x \in X$
$P_2 : (x,y) \in X \times Y \longmapsto y \in Y$

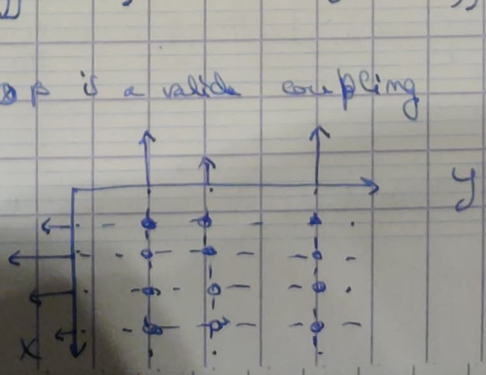
$\pi_1 := (P_1)_\# \pi$ (1st marginal)
$\pi_2 := (P_2)_\# \pi$

Proposition : ① $\Leftrightarrow \pi_1 = \alpha$
② $\Leftrightarrow \pi_2 = \beta$

Example : $\alpha \otimes \beta =$ "$\alpha(x)\beta(y)$"

$$\iint h(x,y)\, d\alpha\otimes\beta(x,y) = \iint h(x,y)\, d\alpha(x)\, d\beta(y)$$

Prop : $\alpha \otimes \beta$ is a valid coupling



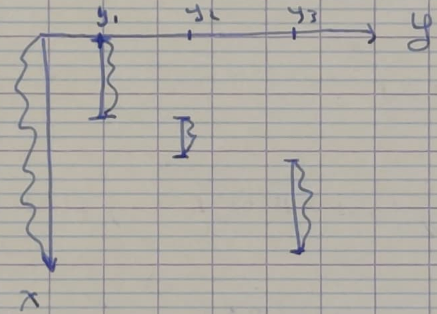
discrete
discrete

Prop: if (*) holds,

$\pi_1 = \alpha$
$\pi_2 = \beta$ $\Leftrightarrow$ $\pi = \sum_{ij} P_{ij} \delta(x_i, y_j)$ $\quad P\mathbb{1} = a, \ P^T\mathbb{1} = b$

$X = \bigcup_{j=1}^{m} \Omega_j$ disjoint $\quad T : x \in \Omega_j \to y_j$ (Piece wise constant)

Prop: iff $\alpha(\Omega_j) = b_j$, $T_\# \alpha = \beta$

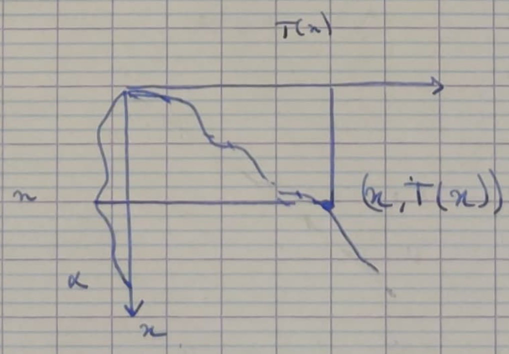


continuous
discrete

Prop: Suppose $T_\# \alpha = \beta$

Then $\pi = (Id, T)_\# \alpha$ is a valid coupling

$\varphi : x \in X \mapsto (x, T(x)) \in X \times Y$



$\boxed{K}$ $\inf_{\pi \in \Gamma(X,Y)} \left\{ \iint c(x,y) \, d\pi(x,y) \ : \ \begin{matrix} \pi_1 = \alpha \\ \pi_2 = \beta \end{matrix} \right\}$

Rmq: if (*) holds (discrete case), $\pi^* = \sum_{ij} P_{ij}^* \delta(x_i, y_j)$

P^* is a sol to discrete kanto.

Prop : if α, β are compactly supported (equiv. $\mathcal{X}$ and $\mathcal{Y}$ are compact)
then K has sol°

Proof : For the weak* topology : • $P(\mathcal{X} \times \mathcal{Y})$ is compact
• $\pi \longmapsto \iint c \, d\pi$ is continuous
• $\alpha \otimes \beta$ is a valid coupling

Prop : For non compact space, $c(x,y) = d(x,y)^p$ $p \geq 1$ $\mathcal{X} = \mathcal{Y}$
finite p-moment $\int d(x_0, x)^p \, d\alpha(x) < +\infty$

Then $\boxed{K}$ has a sol°

Thm : [Brenier] $\mathcal{X} = \mathcal{Y} = \mathbb{R}^d$ $c(x,y) = \|x-y\|^2$
Hyp : α has density / Lebesgue
$\exists !$ sol° $T = \nabla \varphi$ of Monge $(\nabla\varphi)_\# \alpha = \beta$
$\pi = (Id, T)_\# \alpha$ is the ! sol° of Kanto $M = K$

$\alpha \in \mathcal{P}(\mathcal{X})$ " $\Leftrightarrow$ " α is the law of a r.v. X | $\int_A d\alpha(x) = \alpha(A) = \mathbb{P}(X \in A)$
| $\int f(x) \, d\alpha(x) = \mathbb{E}[f(X)]$

$\pi \in \mathcal{P}(\mathcal{X} \times \mathcal{Y}) \Leftrightarrow \pi$ is the law (X, Y)
$\pi_1 = \alpha \Leftrightarrow X$ has law α

$$\boxed{K} \Leftrightarrow \inf_{(X,Y)} \left\{ \mathbb{E}[c(X,Y)] : \begin{array}{l} \text{law}(X) = \alpha \\ \text{law}(Y) = \beta \end{array} \right\}$$

$\pi = \alpha \otimes \beta \Leftrightarrow X$ & Y are independant
$\pi = (Id, T)_\# \alpha \Leftrightarrow Y = T(X)$ "total dependancy"

Def : $X = Y$, $c(x,y) = d(x,y)^p$ $p \geq 1$

Wasserstein p distance $W_p(\alpha,\beta) := \left[\min_{\pi \in \Pi(\alpha,\beta)} \iint d(x,y)^p \, d\pi(x,y) \right]^{1/p}$

Thm : $(\mathcal{P}_p(X), W_p)$ is a metric space

$W_p(\alpha,\beta) = 0 \iff \alpha = \beta$

$(\alpha,\beta,\gamma) \in \mathcal{M}(X)^3$, $W_p(\alpha,\gamma) \leq W_p(\alpha,\beta) + W_p(\beta,\gamma)$

$$\begin{cases} \alpha = \sum a_i \delta_{x_i} \\ \beta = \sum b_j \delta_{y_j} \\ \gamma = \sum c_k \delta_{z_k} \end{cases} \qquad \sum a_i = \sum b_j = \sum c_k = 1$$

$\sum d_{ij}^p P_{ij}^* = 0 \iff P = \mathrm{diag}(h)$

$P\mathbb{1} = a = h$

$P^T\mathbb{1} = b = h$ $\Rightarrow a = b$

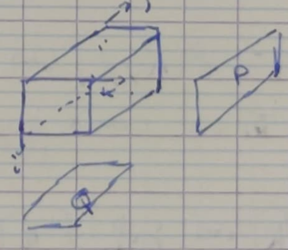
$$a \xrightarrow{P} b \xrightarrow{Q} c$$

$P\mathbb{1} = a$ $Q\mathbb{1} = b$

$P^T\mathbb{1} = b$ $Q^T\mathbb{1} = c$

Lemma : (Gluing)

$G_{ijk} = \frac{P_{ij} Q_{jk}}{b_j}$



$$\begin{cases} \sum_k G_{ijk} = P_{ij} \\ \sum_i G_{ijk} = Q_{jk} \end{cases} \qquad \begin{cases} \sum_{jk} G_{ijk} = a_i \\ \sum_{ik} G_{ijk} = b_j \\ \sum_{ij} G_{ijk} = c_k \end{cases}$$

Def / G : $R_{ik} = \sum_j G_{ijk} = \sum_j \frac{P_{ij} Q_{jk}}{b_j} \iff R = P \,\mathrm{diag}\left(\frac{1}{b}\right) Q$

one has $R\mathbb{1} = b$, $R^T\mathbb{1} = c$

Take P & Q to be optimal $\rightarrow$ R coupling btw a & c

(suboptimality)

$$W_p(\alpha, \beta)^p \leq \left[\sum_{ik} D_{ik}^p \times \sum_j \frac{P_{ij} Q_{jk}}{b_j}\right]^{\frac{1}{p}}$$

$$\leq \left[\sum_{ijk} (D_{ij} + D_{jk})^p \frac{P_{ij} Q_{jk}}{b_j}\right]^{\frac{1}{p}}$$

Minkowsky $\left(\sum_s d_s (a_s + b_s)^p\right)^{\frac{1}{p}} \leq \left(\sum d_s a_s^p\right)^{\frac{1}{p}} + \left(\sum d_s b_s^p\right)^{\frac{1}{p}}$

$p \geq 1$

Minkowsky

$$W_p(\alpha, \beta) \leq \underbrace{\left(\sum_{ijk} D_{ij}^p \frac{P_{ij} Q_{jk}}{b_j}\right)^{\frac{1}{p}}}_{} + \underbrace{\left(\sum_{ijk} D_{jk}^p \frac{P_{ij} Q_{jk}}{b_j}\right)^{\frac{1}{p}}}_{\sum_{ij} D_{ij}^p Q_{jk}}$$

$$= \sum_{ij} D_{ij}^p \underbrace{\sum_k G_{ijk}}_{= P_{ij}}$$

optimality

$$= W_p(\alpha, \beta) + W_p(\beta, \gamma)$$

Thm: $\left.\begin{matrix}\pi \in \mathcal{P}(X \times X) \\ \xi \in \mathcal{P}(X \times X)\end{matrix}\right\} \exists \rho \in \mathcal{P}(X \times X \times X)$

$\rho_{12} = (P_{12})_{\#} \rho = \pi$ $\qquad P_{12} : (x, y, z) \mapsto (x, y)$

$\rho_{23} = (P_{23})_{\#} \rho = \xi$ $\qquad P_{23} : (x, y, z) \mapsto (y, z)$

$\qquad P_{13} : (x, y, z) \mapsto (x, z)$

$$\underbrace{\alpha \overset{\pi}{\longleftrightarrow} \beta \overset{\xi}{\longleftrightarrow} \gamma}_{(P_{13})_{\#} \rho}$$

Total variation: $\|\alpha - \beta\|_{TV} = |\alpha - \beta|(X)$

$\mu = \sum_i u_i \delta_{x_i}$ $\qquad |\mu| = \sum |u_i| \delta_{x_i}$

$\frac{d\mu}{dx} = \rho_\mu$ $\qquad \rho_{|\mu|} = |\rho_\mu|$

EX1: $\alpha = \sum_{i=1}^n a_i \delta_{x_i}$ $\qquad \beta = \sum_{i=1}^n b_i \delta_{x_i}$

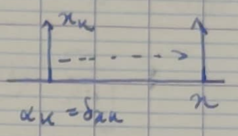
$\|\alpha - \beta\|_{TV} = \|a - b\|_{\ell_1} = \sum_i |a_i - b_i|$

If $d\alpha = \rho_\alpha \, dx \qquad d\beta = \rho_\beta \, dx$

$\|\alpha - \beta\|_{TV} = \|\rho_\alpha - \rho_\beta\|_{L^1(dx)} = \int |\rho_\alpha(x) - \rho_\beta(x)| \, dx$

$(\alpha_k)_{k \geq 0} \xrightarrow{k \to +\infty} \alpha$

$\alpha_k \in \mathcal{P}(X)$



$x_k \neq x \quad \|\delta_{x_k} - \delta_x\|_{TV} = 2 \not\to 0$

$W_p(\delta_{x_k}, \delta_x) = d(x_k, x) \xrightarrow[x_k \to x]{} 0$

Def : $\alpha_k \xrightarrow{\text{strong}} \alpha \iff \|\alpha - \alpha_k\|_{TV} \to 0$

Def : weak-Topology $\iff \left[\forall f \in \mathcal{C}(X), \int f(x) \, d\alpha_k(x) \to \int f(x) \, d\alpha(x)\right]$

$\alpha_k \xrightarrow{*} \alpha$

On X compact $\mathcal{M}(X) = \mathcal{C}(X)^*$

$\|\cdot\|_{TV} \qquad \|f\|_\infty$

Proposition : $\|\alpha\|_{TV} = \sup_{\|f\|_\infty \leq 1} \int f(x) \, d\alpha(x)$

" $\langle f, \alpha \rangle$ "

Rmq : $(B, \|\cdot\|)$

$(B^*, \|\cdot\|_*)$ where $\|\cdot\|_* = \sup_{\|x\| \leq 1} \langle x, y \rangle$

⚠ on non ~~compact~~ compact space, $\mathcal{M}(X) = \mathcal{C}_0(X)^*$

$\mathcal{C}_0(X) =$ " $f, \ f(x) \xrightarrow[x \to +\infty]{} 0$ "

$= \{f, \forall \varepsilon, \exists K \text{ Compact}, x \notin K, |f(x)| \leq \varepsilon\}$

Rmq : $\alpha_k \sim X_k \qquad \alpha_k \xrightarrow{*} \alpha \iff X_k \xrightarrow{\mathcal{L}} X$

$\alpha \sim X$

$\iff \forall f \in \mathcal{C}_b(X), \ \mathbb{E}[f(X_k)] \to \mathbb{E}[f(X)]$

Loi des grands nombres :

$\mathbb{E}[\|X\|] < +\infty$

$$\frac{X_1 + \dots + X_n}{n} \xrightarrow[\mathcal{L}]{a.s.} \mathbb{E}[X]$$

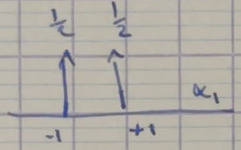
Theorem: $X_1, \dots X_n$ iid $\mathbb{E}[X] = 0$ $\mathbb{E}[X_i^T X_i] = C$

$\mathbb{E}[|X|^2] < +\infty$

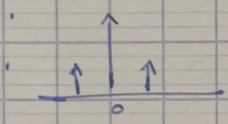
$$Y_n = \frac{X_1 + \dots + X_n}{n}$$

$$\sqrt{n}\, Y_n \xrightarrow{\mathcal{L}} \mathcal{N}(0, C)$$

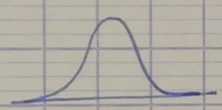
$$\Updownarrow \quad \alpha_n \xrightarrow{*} \mathcal{N}(0, C)$$



X_1



α_2 $\quad X_1 + X_2$ $\quad$ Convolution $\alpha_2 = \alpha_1 * \alpha_1$



$$\|\alpha_n - \mathcal{N}(0,1)\|_{TV} = 2$$

Theorem: $\mathcal{X}$ compact $\alpha_n \xrightarrow{*} \alpha \iff W_p(\alpha_n, \alpha) \xrightarrow{n \to \infty} 0$

In general: $W_p(\alpha_n, \alpha) \to 0 \iff \begin{cases} \alpha_n \xrightarrow{*} \alpha \\ \int d(x_0, x)^p \, d\alpha_n \to \int d(x_0, x)^p \, d\alpha \end{cases}$

Thm: If $\mathbb{E}[\|X\|^3] < +\infty$, then $\exists C$

$$W_p(\alpha_n, \mathcal{N}(0,1)) \leq \frac{C}{\sqrt{n}}$$